LiNets: Locally-Connected, Irregular Deep Neural Networks for Biomimetic Active Vision in a Simulated Human

Masaki Nakada, Honglin Chen, Arjun Lakshmipathy, and Demetri Terzopoulos
(nakada@cs.ucla.edu)

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

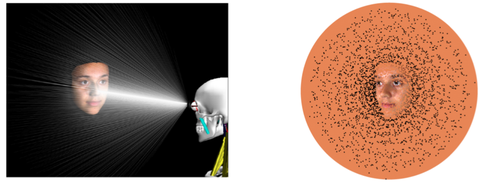
Overview:

- We introduce locally-connected, irregular deep neural networks (liNets) for biomimetic active vision. Like common CNNs, liNets are locally-connected, forming receptive fields, but unlike CNNs, they are suitable for spatially irregular photoreceptor distributions like those found in foveated biological retinas.
- Compared to fully-connected deep neural networks, liNets accommodate a much greater number of retinal photoreceptors to enhance visual acuity without intractable memory consumption.
- LiNets serve well in the biomimetic active vision system embodied in a simulated human that learns active visuomotor control and active appearance-based recognition.

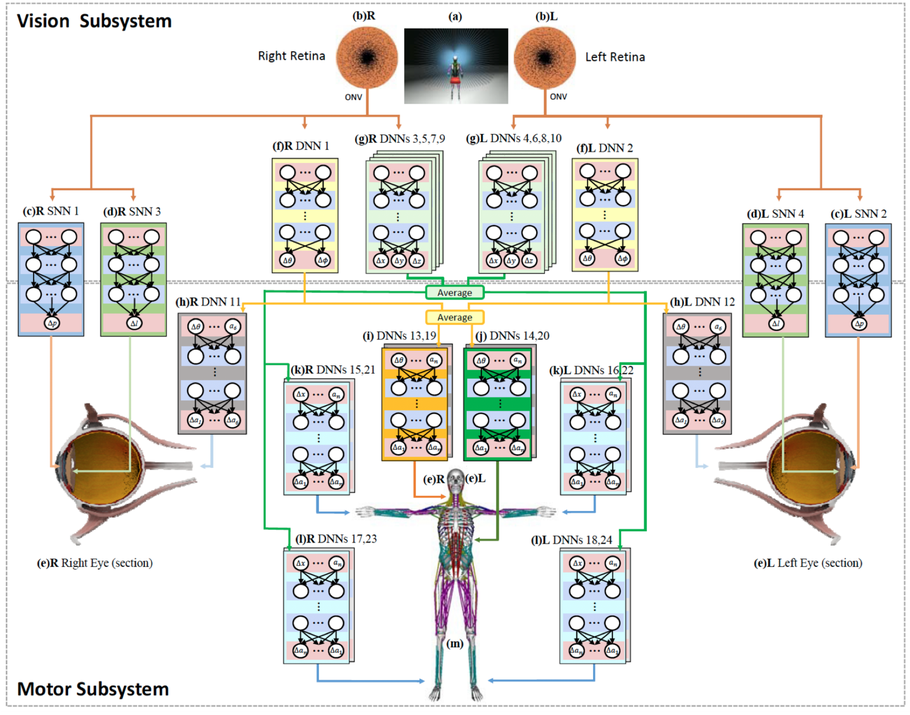
LiNets:

- In a fully-connected neural network, each neuron in a hidden layer is connected globally to all the neurons in the previous layer, whereas like CNNs, each neuronal unit in a hidden layer of our liNet is connected only locally to a fixed number of neighboring units in the previous layer. However, the implementation of a CNN conforms to conventional, regularly sampled images, the regular arrangement of units in tensor data structures, and the sharing of weights and biases across all the units in each hidden layer (i.e., the convolutional property). The liNet shares none of these restrictions. Importantly, every unit of a liNet has its own particular weights and bias, just like in a fully-connected network.

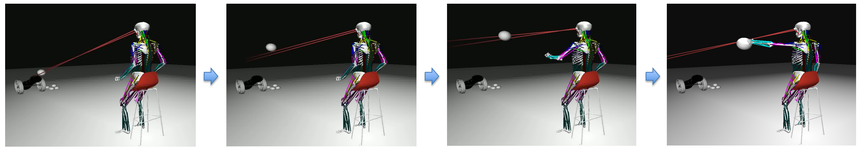
Active 3D face recognition



Sensorimotor system architecture

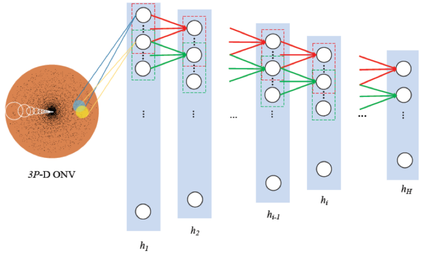


The biomechanical virtual human actively executing arm reaching movements to intercept balls



Photoreceptor responses in a saccadic eye movement to foveate a white ball

The liNet architecture



- Suppose we have a liNet with H hidden layers, each of which is comprised of N_h neuronal units. Using a Euclidean-distance k-nearest-neighbor algorithm, each unit in hidden layer h = 1,…,H forms a receptive field with R_h nearest neighbor units in the previous hidden layer h-1, and it inherits a position in the 2D (polar coordinate system) retinal domain that is the average of the positions of the units in its receptive field. The receptive fields of the units in the first hidden layer h = 1, are formed from R_1 neighboring retinal photoreceptors. Given an irregular, foveated photoreceptor distribution, the overlapping receptive fields, which are illustrated by the white circles in the retinal domain at the left of the figure naturally increase in size with eccentricity, from the denser foveal center outward to the sparser periphery.

Sensorimotor System:

- Sensorimotor system architecture (left), showing the modular neural network controllers in the sensory and motor subsystems, including a total of 24 DNNs.
- ***Vision Subsystem:*** To compute irradiance via ray-tracing, each retinal photoreceptor casts rays through the eye (Fig. 2a,b,c) and into the virtual world (a). (b) The arrangement of the photoreceptors (black dots) on the right R and left L foveated retinas. Each eye outputs an RGB Optic Nerve Vector (ONV). This feeds four trained visual accommodation SNNs (1–4); SNNs (c)R (1) and (d)R (3) control the muscles of the iris and lens of the right eye (e)R, and SNNs (c)L (2) and (d)L (4) do the same for the left eye (e)L. The ONV also feeds ten trained vision DNNs (1–10). (f) A pair of foveation DNNs (1,2) produce outputs that drive the movements of the eyes to foveate visual targets. (g) The eight limb vision DNNs (3–10) — (g)R (3,5,7,9) for the right eye and (g)R (4,6,8,10) for the left eye — output observed limb-to-target discrepancy estimates.
- ***Motor Subsystem:*** Fourteen trained neuromuscular motor DNNs (11–24) comprise the motor subsystem, including eight voluntary motor DNNs (11–18) and six reflex motor DNNs (19–24). (h) The oculomotor DNNs (11,12), which are driven by the outputs of the foveation DNNs, output muscle activation signals that control the six extraocular muscles of each eye to produce eye movements. Driven by the averaged responses of the foveation DNNs, along with the current activations of the 216 neck muscles and 443 torso muscles, respectively, the cervicocephalic (i) voluntary motor DNN (13) and torso (j) voluntary motor DNN (14) each outputs muscle activation signals that contribute to actuating its associated neuromuscular complex. Driven by the bilaterally pairwise averaged responses of the limb vision DNNs, along with the current activations of the 29 muscles of each arm or 39 muscles of each leg, respectively, each of the four limb voluntary motor DNNs (k) (l) (15–18) outputs muscle activation signals that contribute to actuating its associated neuromuscular complex. Each of the six reflex motor DNNs (19–24) outputs muscle activation signals that contribute by stabilizing the muscle group of its associated musculoskeletal complex.